UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/314,637	05/19/1999	MAZIN G. RAHIM	ATT.0030000	5478

7590 06/17/2005

EXAMINER
ARMSTRONG, ANGELA A

Mr. S H Dworetsky
AT&T Corp
P O Box 4110
Middletown, NJ 07748

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 06/17/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/314,637	RAHIM	
	Examiner	Art Unit	
	Angela A. Armstrong	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *23 May 2005*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *13-27 and 29-36* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *13-27 and 29-36* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on May 23, 2005, has been entered.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 17-19, 21-27, 29-34, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Alleva et al (US Patent No. 5,970,449) in view of Sukkar (US Patent No. 5,613,037).

3. Regarding claim 17, Alleva at al teaches

receiving a speech signal at col. 3, line 18 to col. 4, line 6;

Alleva describes the speech recognition processor that produces textual output corresponding to recognized portions of input speech, such that the recognizer produces text such as "ten cents" and "four o'clock in the afternoon," which reads on "performing speech recognition process on the received speech signal to produce speech recognition results, wherein

a numeric language includes a subset of a vocabulary, the subset of the vocabulary including words that identify digits in number strings and words that enable the interpretation and understanding of number strings," since the words ten, cents, four and o'clock are the vocabulary words of numeric language that are relevant for interpreting and understanding number strings related to currency and time (col. 3, line 18 to col. 4, line 6; Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56-62; col. 6, lines 14-17 and 40-42; col. 5, lines 62-65 and col. 6, lines 32-64);

At col. 6, lines 14-64 and Figure 9, elements 122, 124, 126, 128, and 130, Alleva et al describes the rules the text normalizer (element 38, Figures 3A-3E) implements to process the string of words received from the speech recognizer to generate a sequence of corresponding digits, which reads on "generating a sequence of digits using said speech recognition results, said generating being based on a set of rules."

Alleva fails to explicitly teach a system comprising acoustic models utilized by the speech recognition processor. However, implementation of acoustic models in a speech recognition system was well known in the art.

In a similar field of endeavor, Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 18, Alleva teaches the speech recognition processor at Figure 1, element 32; col. 3, line 18 to col. 4, line 6.

Regarding claim 19, Alleva does not teach that the recognition process on a set of acoustical models that has been defined for other words in the vocabulary.

Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 21, Alleva teaches the speech recognition processor that produces textual output corresponding to recognized portions of input speech, such that the recognizer produces text such as "ten cents," "April first nineteen ninety seven," "Seattle Washington nine eight zero five two" and "four o'clock in the afternoon," which reads on "numeric language includes digits, natural numbers, alphabets, and city/country name classes," since the words ten, cents, April, Seattle, Washington, four and o'clock are the vocabulary words of numeric language that are relevant for interpreting and understanding number strings related to classes of digits, natural numbers, alphabets, and city/country name (col. 3, line 18 to col. 4, line 6;

Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56-62; col. 6, lines 14-17 and 40-42; col. 5, lines 62-65 and col. 6, lines 32-64).

Alleva does not teach the numeric language includes a re-starts class. At col. 5, line 48-52, Sukkar discloses implementation of a misrecognition classifier, so as to account for the errors during recognition.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement words in the numeric language related to recognition errors to account for errors during the recognition process, as suggested by Sukkar, for the purpose of providing reliable and accurate recognition and thereby improve system performance.

Regarding claim 22, Alleva does not explicitly teach acoustic models are hidden Markov models. Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 23, Alleva does not teach a numeric recognition processor. Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement a numeric recognition processor as taught by Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of recognized numbers of the received input speech.

Regarding claims 24 and 26-27, Alleva teaches a set of rules includes a naturals rule, a restarts rule, a city/country rule, and a alphabets rule at Figure 9, element 126 and col. 6, line 3 to col. 7, line 9.

Regarding claim 25, Alleva does not teach the set of rules includes re-starts rules. At col. 5, line 48-52, Sukkar discloses implementation of a misrecognition classifier, so as to account for the errors during recognition.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement process or normalize the words in the numeric language output from the speech recognizer that are related to recognition errors to account for errors during the recognition process, as suggested by Sukkar, for the purpose of providing reliable and accurate recognition and thereby improve system performance.

Regarding claim 29, Alleva et al teaches a system for text normalization in which the output of a speech recognizer is processed to produce a representation of the appropriate digits. Alleva describes the speech recognition processor that produces textual output corresponding to recognized portions of input speech, such that the recognizer produces text such as "ten cents" and "four o'clock in the afternoon", which reads on "a speech recognition processor that receives unconstrained input speech and outputs a string of words, the speech recognition processor being based on a numeric language that represents a subset of a vocabulary, the subset including a set

of words identified as being relevant for interpreting and understanding number strings," since the words ten, cents, four and o'clock are the vocabulary words of numeric language that are relevant for interpreting and understanding number strings related to currency and time (col. 3, line 18 to col. 4, line 6; Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56-62; col. 6, lines 14-17 and 40-42; col. 5, lines 62-65 and col. 6, lines 32-64);

At col. 6, lines 14-64, Alleva et al describes the rules the text normalizer (element 38, Figures 3A-3E) implements to process the string of words received from the speech recognizer to generate a sequence of corresponding digits, which reads on "a numeric understanding processor containing classes of rules for converting the string of words into a sequence of digits."

Alleva fails to explicitly teach a system comprising acoustic models utilized by the speech recognition processor. However, implementation of acoustic models in a speech recognition system was well known in the art.

Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 30, Alleva fails to explicitly teach a first set of hidden Markov models that characterize acoustic features of words in the numeric language and a second set of hidden Markov models that characterize acoustic features of words in the remainder of the vocabulary.

Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic hidden Markov model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 31, Alleva fails to explicitly teach a filler model that characterizes out of vocabulary features.

Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic hidden Markov model teachings of a filler model, as suggested by Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech to accurately distinguish numeric input from other speech input.

Regarding claim 32, Alleva fails to teach an utterance verification processor. At col. 5, lines 44-52, Sukkar describes a digit/non-digit classification that identifies speech containing valid digits, speech not containing a digit and speech containing misrecognitions. Sukkar teaches the misrecognitions are identified as non-digits so that errors can be rejected and not classified as valid digit data.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Alleva to implement utterance verification as taught by Sukkar, for the purpose of ensuring that only valid digit information is recognized and classified as actual digit speech.

Regarding claim 33, Alleva does not teach a validation database or a string validation processor. At col. 7, lines 6-49, Sukkar describes candidate string validation based on individual candidate digit confidence scores that are determined using a digit vocabulary set of the digit models. Sukkar teaches the string validation is implemented so that errors in the string cause the string to be rejected, which is desirable for many applications.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Alleva to implement string validation as taught by Sukkar, for the purpose of ensuring that only valid digit information is accepted and applications using the system process and operate with valid data.

Regarding claim 34, at col. 8, lines 28-38, Alleva teaches the normalizer normalizes the text and a speech API forwards the normalized content to a application program, which reads on "a dialogue manager processor that initiates an action based on the validity information."

Regarding claim 36, Alleva teaches a set of rules includes a naturals rule, a restarts rule, a city/country rule, and a alphabets rule at Figure 9, element 126 and col. 6, line 3 to col. 7, line 9. Alleva does not teach the set of rules includes a re-starts rule. At col. 5, line 48-52, Sukkar discloses implementation of a misrecognition classifier, so as to account for the errors during recognition.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement process or normalize the words in the numeric language output from the speech recognizer that are related to recognition errors to account for errors during the recognition process, as suggested by Sukkar, for the purpose of providing reliable and accurate recognition and thereby improve system performance.

4. Claims 13-16, 20 and 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Alleva et al (US Patent No. 5,970,449) in view of Sukkar (US Patent No. 5,613,037) and further in view of Huang et al (US Patent No. 5,937,384).

5. Regarding claim 13, Alleva describes the speech recognition processor that produces textual output corresponding to recognized portions of input speech, such that the recognizer produces text such as "ten cents" and "four o'clock in the afternoon," which reads on "a speech recognition method comprising, defining a numeric language, the numeric language including a subset of a vocabulary, the subset of the vocabulary including words that identify digits in number strings and words that enable the interpretation and understanding of number strings," since the words ten, cents, four and o'clock are the vocabulary words of numeric language that are relevant for interpreting and understanding number strings related to currency and time (col. 3, line 18 to col. 4, line 6; Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56-62; col. 6, lines 14-17 and 40-42; col. 5, lines 62-65 and col. 6, lines 32-64);

Alleva does not teach a set of acoustic models for the numeric language, a second set of acoustical models that has been defined for other words in the vocabulary or storing the first and second set of acoustical models in an acoustic model database.

Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Alleva and Sukkar do not implement a first quality level for the first acoustic models and a second quality level for the second acoustic models. Huang teaches a method and system for speech recognition using continuous density hidden Markov models, which implements context-dependent HMMs and context-independent HMMs and teaches that the use of both types of HMMs is beneficial in achieving an improved recognition accuracy (col. 6, lines 18-38).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Alleva and Sukkar to implement both context-dependent HMMs and context-independent HMMs, as by Huang, for the purpose of achieving an improved recognition accuracy, as suggest by Huang.

Regarding claim 14, Alleva teaches the speech recognition processor that produces textual output corresponding to recognized portions of input speech, such that the recognizer produces text such as "ten cents," "April first nineteen ninety seven," "Seattle Washington nine eight zero five two" and "four o'clock in the afternoon," which reads on "numeric language includes digits, natural numbers, alphabets, and city/country name classes," since the words ten, cents, April, Seattle, Washington, four and o'clock are the vocabulary words of numeric

language that are relevant for interpreting and understanding number strings related to classes of digits, natural numbers, alphabets, and city/country name (col. 3, line 18 to col. 4, line 6; Abstract; Figure 1, element 32; Figure 9, element 132; col. 1, lines 56-62; col. 6, lines 14-17 and 40-42; col. 5, lines 62-65 and col. 6, lines 32-64).

Alleva does not teach the numeric language includes a re-starts class. At col. 5, line 48-52, Sukkar discloses implementation of a misrecognition classifier, so as to account for the errors during recognition.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement words in the numeric language related to recognition errors to account for errors during the recognition process, as suggested by Sukkar, for the purpose of providing reliable and accurate recognition and thereby improve system performance.

Regarding claim 15, Alleva does not explicitly teach acoustic models are hidden Markov models. Sukkar discloses a speech recognition system comprising acoustic model, utilized by the speech recognition processor (Figure 3, element 308). Additionally, Sukkar teaches a digit model for digit recognition and a second model, a filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic model teachings of Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech.

Regarding claim 16, Alleva fails to explicitly teach a filler model that characterizes out of vocabulary features. Sukkar teaches a digit model for digit recognition and a second model, a

filler model, a generalized HMM model of spoken words that do not contain digits (col. 3, line 19 to col. 4, line 22).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to implement the acoustic hidden Markov model teachings of a filler model, as suggested by Sukkar in the recognition system of Alleva, for the purpose of accurately producing vector representations of the received input speech to accurately distinguish numeric input from other speech input.

Regarding claim 20, Alleva and Sukkar do not implement a first quality level for the first acoustic models and a second quality level for the second acoustic models. Huang teaches a method and system for speech recognition using continuous density hidden Markov models, which implements context-dependent HMMs and context-independent HMMs and teaches that the use of both types of HMMs is beneficial in achieving an improved recognition accuracy (col. 6, lines 18-38).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Alleva and Sukkar to implement both context-dependent HMMs and context-independent HMMs, as by Huang, for the purpose of achieving an improved recognition accuracy, as suggest by Huang.

Regarding claim 35, Alleva and Sukkar do not specifically teach a language model database that stores data describing the structure and sequence of words and phrases. Huang teaches a language model that represents linguistic expressions and describes the implementation of language model in predicting the likelihood of occurrence of a word considering the words

that have been uttered (col. 14, lines 35-54) and teaches the system is beneficial in improving the recognition capability of a speech recognition system.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Alleva and Sukkar to implement language models in predicting likelihoods of word occurrences, as taught by Huang, for the purpose of improving recognition capability of the speech recognizer.

Response to Arguments

6. Applicant's arguments filed May 23, 2005, have been fully considered but they are not persuasive.

In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988)and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, the Examiner contends implementation of acoustic models representative of input speech to be recognized for implementation in a speech recognition system was well known to one of ordinary skill in the art of speech signal processing to achieve improved accuracy. The Examiner argues that implementation of the digit/filler/generalized HMM models of Sukkar improves the vector representations of the received input speech digits and thereby improve the

speech recognition accuracy of the recognizer. Therefore, the Examiner maintains, there is sufficient motivation for the combination of the teachings of Alleva and Sukkar, and thus the rejection is maintained.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela A. Armstrong whose telephone number is 571-272-7598. The examiner can normally be reached on Monday-Thursday 11:30-8:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Angela A Armstrong
Examiner
Art Unit 2654

AAA
June 12, 2005

Angela Armstrong